·智能网联汽车安全验证与评估专题·

自动驾驶系统部署后安全治理中的在用监测与报告机制研究*

陈贞[1]　杨强[2]　陈俊涛[3]

（1.北京镝石数据科技有限公司，北京 100176；2.北京赛目科技股份有限公司，北京 100080；3.北京航迹科技有限公司，北京 100080）

【欢迎引用】陈贞，杨强，陈俊涛．自动驾驶系统部署后安全治理中的在用监测与报告机制研究[J]．汽车文摘，2026(8)：27-36.

【Cite this paper】CHEN Z, YANG Q, CHEN J T. Research on In-Service Monitoring and Reporting for Post-Deployment Safety Governance of Automated Driving Systems[J]. Automotive Digest (Chinese), 2026(8): 27-36.

【摘要】针对自动驾驶系统（ADS）软件持续迭代、运行环境复杂及AI模型动态演化等特征导致传统静态型式批准模式难以适配全生命周期安全治理需求的挑战，基于制度分析与试点实践，构建面向ADS部署后安全治理的在用监测与报告（ISMR）机制及其实现路径。采用规范分析与机制设计方法，提出了覆盖运行数据获取、数据治理、安全分析、监测报告和安全改进的企业侧实施机制；构建了“规则供给-运行执行-证据反馈”的政企协同治理框架；设计了事件触发的部署后安全评估框架。研究表明，该机制以运行证据为核心，能够有效支撑ADS的持续安全保证，实现对运行风险的持续识别、分析与闭环处置，并为安全档案的动态更新提供依据。

关键词：自动驾驶系统；在用监测与报告（ISMR）；部署后安全评估；全生命周期安全治理；协同治理

中图分类号：U461.91　　**文献标志码：**A　　**DOI:** 10.19822/j.cnki.1671-6329.20260083

Research on In-Service Monitoring and Reporting for Post-Deployment Safety Governance of Automated Driving Systems

Chen Zhen[1], Yang Qiang[2], Chen Juntao[3]

（1. Beijing Dishi Data Technology Co., Ltd., Beijing 100176; 2. Beijing Saimo Technology Co., Ltd., Beijing 100080; 3. Beijing Voyager Technology Co., Ltd., Beijing 100080）

【Abstract】To address the challenge that the traditional static type approval model struggles to accommodate the full lifecycle safety governance of Automated Driving Systems (ADS), which are characterized by continuous software iteration, complex operational environments, and dynamic evolution of AI models, the paper constructs an In-Service Monitoring and Reporting (ISMR) mechanism and its implementation pathway based on institutional analysis and pilot practices. Employing normative analysis and mechanism design methods, a enterprise-side implementation mechanism covering operational data acquisition, data governance, safety analysis, monitoring reporting and safety improvement is proposed. Furthermore, a government-enterprise collaborative governance framework featuring “rule supply, operational execution, and evidence feedback” is established, alongside an event-triggered post-deployment safety assessment framework. The results demonstrate that the proposed mechanism, centered on operational evidence, effectively supports the continuous safety assurance of ADS. It enables the ongoing identification, analysis, and closed-loop mitigation of operational risks while providing a basis for the dynamic updating of safety cases.

Key words: Automated Driving Systems (ADS), In-Service Monitoring and Reporting (ISMR), Post-deployment safety assessment, Lifecycle safety governance, Collaborative governance

*基金项目：新一代人工智能国家科技重大专项（2022ZD0116311）。

0 引言

自动驾驶系统(Automated Driving System, ADS)具有软件持续迭代、运行环境开放复杂和AI模型行为动态演化等特征[1-2]。传统基于定型样车与有限工况验证的静态型式批准模式,难以覆盖AI不确定性、长尾场景风险及系统部署后持续演化带来的新增风险,导致准入阶段的安全承诺在运行阶段难以持续保持和验证[3-5]。

为应对上述挑战,联合国世界车辆法规协调论坛(UNECE WP.29)自动驾驶系统全球技术法规(GTR ADS)框架[6]、我国《智能网联汽车准入和上路通行试点实施指南(试行)》等政策法规[7]以及ISO/TS 5083:2025等国际标准[8],均提出在用监测与报告(In-Service Monitoring and Reporting, ISMR)要求,强调对已部署ADS运行表现的持续监测及安全相关事件与绩效指标的报告,用于支撑运行阶段安全状态评估与产品生产一致性验证[3,6-7]。

在实践层面,欧盟L3Pilot试点项目[9]、Waymo等企业应用实践[2]以及我国智能网联汽车准入和上路通行试点[7],均已基于运行数据开展自动驾驶系统部署后安全管理探索,初步形成了以运行证据支撑安全评估与改进的技术路径。

现有研究主要集中于自动驾驶技术研发、安全测试与评价及场景构建等领域,针对部署后运行阶段的系统化治理研究仍较为有限,尤其是在ISMR工程化实施路径、基于运行证据的监管决策机制,以及政企协同治理结构方面,尚缺乏系统研究[10-12]。此外,产业界既有数据闭环机制主要服务于模型训练与能力迭代[14-15],与面向安全保证的制度化监测体系仍存在明显功能区分[5,13]。

针对上述问题,本文融合制度机制与系统工程视角,围绕ISMR在ADS部署后安全治理中的作用机制,从制度逻辑、企业实施机制、协同治理机制、部署后安全评估框架4个层面展开研究,研究框架如图1所示。

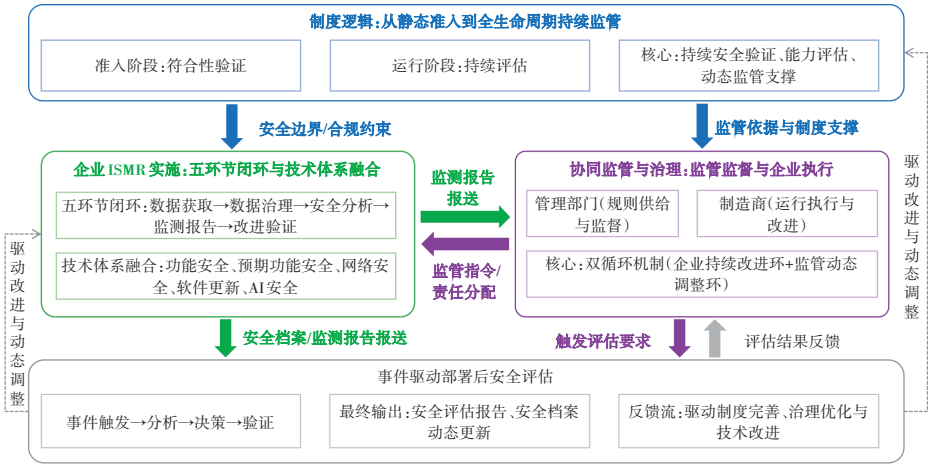


图1 研究框架

1 ADS型式批准生命周期化演进

1.1 基于运行证据的型式批准生命周期演化机制

ISMR作为ADS部署后的安全保障机制,有力支撑传统静态型式批准模式向全生命周期监管模式演进。其核心特征是形成"准入阶段符合性验证与运行阶段持续评估相结合"的双阶段结构(图2)。其中,准入阶段通过测试验证、安全论证及能力审核,形成安全档案与安全承诺;车辆部署后,运行阶段通过ISMR持续获取运行证据,对系统实际运行表现进行跟踪评估,并验证其是否持续满足准入阶段建立的安全档案与风险接受标准。二者形成基于运行证据的闭环验证过程[16]。同时,ISMR与仿真测试、场地测试、道路测试、安全保障能力审核评估共同构成ADS全生命周期安全保证的多支柱验证体系,用于覆盖运行阶段难以在设计开发阶段充分暴露的长尾场景风险与系统动态演化风险[6-8]。

在上述生命周期监管框架下,准入阶段与运行阶段通过运行证据驱动形成持续安全评估机制,该机制呈现为由运行证据生成、汇聚与反馈构成的闭环结构。

(1)准入阶段:ISMR相关能力作为型式批准符合性的重要组成部分纳入验证框架,重点验证自动驾驶数据记录系统(Data Storage System for Automated Driving, DSSAD)、事件数据记录系统(Event Data Recorder, EDR)及ISMR相关的事件触发、数据记录、报告生成等能力的可实现性与合规性[6-7]。同时,对制

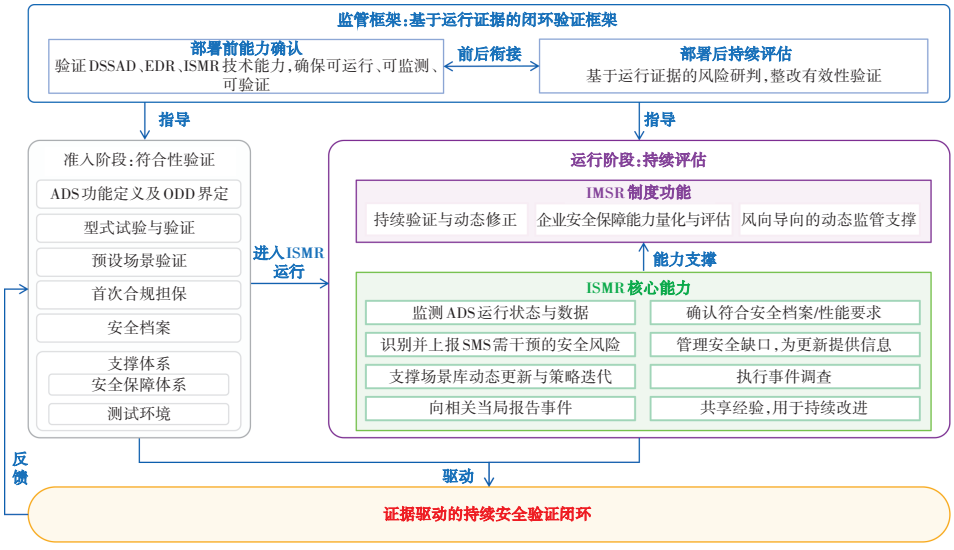


图2　ADS型式批准生命周期化结构与ISMR作用机制

造商在运行阶段支撑ISMR持续运行所需的组织与管理能力进行评估[17]，包括监测流程设计、数据治理机制、工具链支撑能力、关键岗位人员能力等[6-7]。

（2）运行阶段：ISMR以运行数据为基础持续生成安全运行证据，并通过事件分析与指标评价形成结构化安全状态表征。在此过程中，企业侧主要围绕证据生成与风险处置开展持续改进活动，实现系统状态的持续维护与纠偏；监管侧则基于事件报告与安全绩效指标，对系统运行安全状态进行评估与分级管理，并触发相应的治理响应与调整机制[6-7]。二者通过证据流的协同交互作用，共同构成运行阶段安全状态的闭环治理机制。

1.2　ISMR与数据闭环的区别

ISMR与产业界通用的数据闭环虽共享运行数据基础设施，但在制度目标与功能边界上存在显著差异（表1）[6,14]。

表1　ISMR与数据闭环的特征对比[6,14]

维度	ISMR	数据闭环
核心目标	安全状态验证与合规支持	模型与系统能力提升
驱动机制	法规与监管要求	产品迭代与工程优化
数据范围	安全相关事件、异常运行数据及关键绩效指标	正常运行数据与异常数据
处理逻辑	监测-分析-报告-处置	数据训练-模型优化-功能迭代
输出结果	安全状态评估结果、风险处置建议、合规证据	模型、算法与功能升级结果

从功能结构看，数据闭环面向技术优化决策，主要输出模型、算法及功能升级结果；ISMR面向安全治理决策，主要输出安全状态评估结果、风险分析结论及合规证据。因此，两者分别对应ADS运行数据体系中的“能力增强链路”与“安全保证链路”。

在技术研究层面，美国国家标准与技术研究院（NIST）指出，已部署AI系统在持续运行监测方面仍存在性能漂移、不确定性行为及环境变化适应性风险等方面的识别与评估不足，且缺乏标准化监测机制支持，并提出涵盖功能、运行、人为因素、安全、合规及社会影响等维度的系统性监测框架[3]，进一步凸显面向运行阶段安全治理的持续监测机制构建需求。

1.3　ISMR的制度功能

基于制度设计与运行目标，ISMR主要承担持续安全验证、企业安全保障能力评估以及动态监管支撑3项功能：

（1）持续验证与动态修正。ISMR通过事件数据采集等技术手段获取事件标记、系统状态及关键行为数据，为运行阶段安全评估提供可追溯证据[6-7]。基于这些运行证据，可以持续验证ADS在OTA升级、设计运行范围（Operational Design Domain，ODD）扩展及运行环境变化条件下，是否仍满足准入阶段建立的安全档案与风险接受标准。

（2）企业安全保障能力量化与评估。通过建立事件监测、风险分析、整改闭环及报告管理等流程，并采用安全绩效指标（Safety Performance Indicators，SPIs）进行量化评价，可客观反映企业持续安全保障能力的有效性，为管理部门开展能力监督与符合性评估提供依据[6-7]。

（3）风险导向的动态监管支撑。ISMR形成的结构化事件报告与聚合化数据，能够为监管部门提供持续、客观的风险信息来源。一方面，可提升事故调查、风险识别和趋势分析能力；另一方面，可基于运行实践发现

现有监管要求、测试场景及评价指标存在的不足，为法规修订、场景扩展和监管机制优化提供实证支撑[6-7]。

2 制造商ISMR实施机制

将ISMR从监管要求转化为企业内部可运营、可持续并具备安全改进能力的管理实践，是制度落地的关键。基于对法规的解构与企业工程实践[2,6]，并参考行业的持续监测框架[10]，本文构建从数据采集到安全验证的五环节实施机制（图3），并剖析其背后所需的认知转型与核心能力。

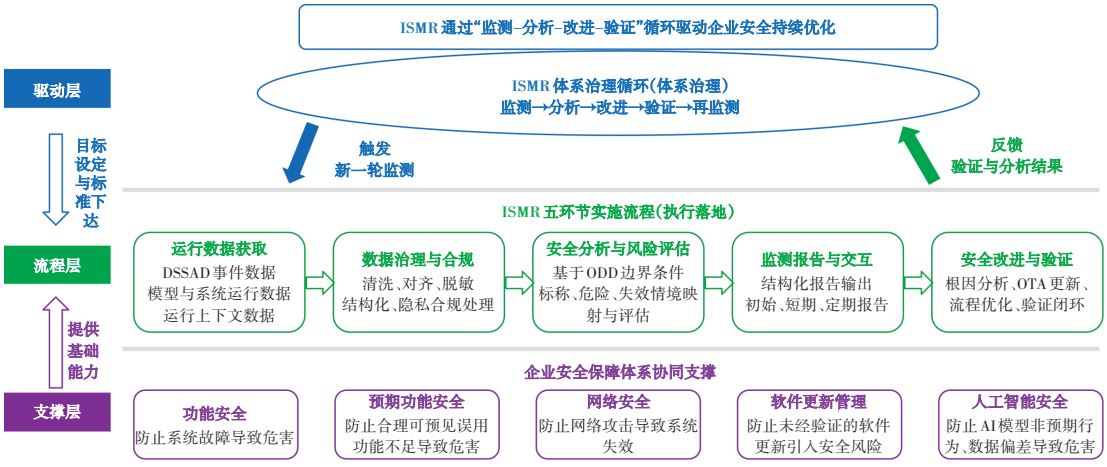


图3 面向持续安全状态维持与治理闭环的ISMR实施框架

2.1 企业责任转变

ISMR制度的实施强化了制造商的全生命周期主体责任[6-7]。制造商责任范围由准入阶段合规延伸至产品运行全过程，需要持续承担运行监测、风险处置和安全改进责任。这需要企业具备数据治理、风险分析、整改与验证等能力，并对ISMR体系设计、运行维护、数据质量及整改落实承担主体责任。

2.2 ISMR实施机制与关键活动

为支撑运行阶段安全风险的持续管理，ISMR需构建覆盖数据获取、数据治理、安全分析、监测报告与整改验证5个环节的实施流程[10]。

2.2.1 运行数据获取

有效的数据获取是ISMR的治理起点。企业需要建立分层数据采集体系，以支撑合规要求与深度分析需求。

（1）法规基准数据。基于DSSAD、EDR等记录事件标记与行为序列数据，保证法定报告事件可触发、可记录、不可篡改，主要包括2类：一是用于事件定性与时序定位的事件标记数据；二是用于还原事件过程的状态标记信息或行为序列数据。此层数据是监管进行风险评估的法定依据。

（2）企业扩展分析数据。为支撑持续安全改进，制造商采集模型运行状态、决策过程与关键环境上下文等扩展数据，可用于事件根因分析、模型运行状态评估，以及整改措施制定。然而，该层数据的获取在技术上受限于模型的可见性。NIST研究指出，当前难以实现模型内部属性（如数据使用、模型组件、模型推理过程）的完全可见性，这在一定程度上增加了运行事件根因分析的难度[3]。

（3）外部反馈数据。为全面解析事件成因，需整合运维反馈、执法信息、用户反馈、交互日志、售后服务等多源信息，为ISMR记录的核心事件分析提供上下文的辅助数据。

2.2.2 数据治理与合规

ISMR数据治理需在满足隐私保护要求的前提下实现数据可用。企业遵循车内处理、最小必要、匿名化等原则，建立满足数据安全与隐私保护要求的数据处理机制[18]。

基于现有实践[9,19]，数据治理可采用“原始数据本地留存、脱敏指标共享”的分层架构（图4），明确企业与管理部门之间的数据管理职责和共享边界[20-21]：原始传感器数据与可识别个人身份信息原则上由企业端存储与管控，仅在满足监管需求且经脱敏或假名化处理后，向管理部门共享事件标记、SPIs及必要的衍生度量值（DMs）。例如，采用加盐哈希（Salting Hash）等数据隐私保护技术对驾驶员身份、行程标识等敏感字段进行处理[21]，兼顾数据安全与分析需求。

同时，企业可通过建立覆盖完整性、准确性、时效性的数据质量评价体系[19]，对事件监测系统记录的触发准确率、时间戳同步偏差等进行常态化监测，并建立数据清洗与补采机制，确保输入至安全分析与监测

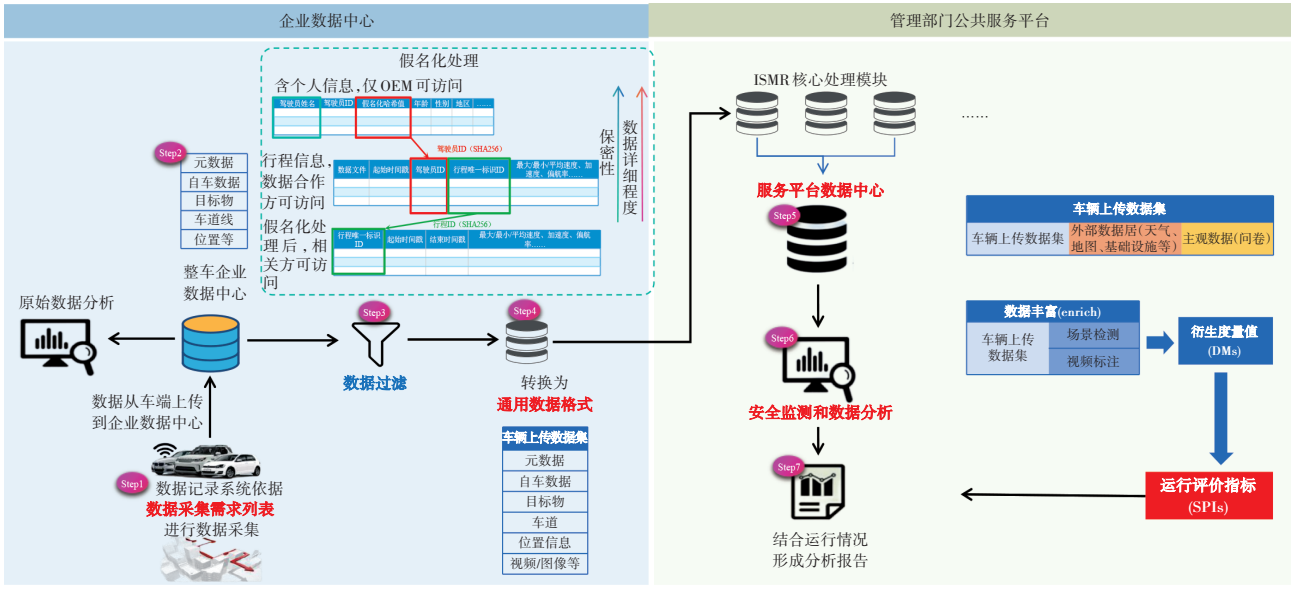


图4　面向ISMR的数据治理与运行评价流程图

报告环节的证据具备工程可用性[22]。

2.2.3　安全分析与风险评估

基于运行数据的安全分析采用个体事件分析与群体运行分析相结合的方式。首先，通过常规参数监测建立系统运行基线；其次，SPIs超限检测，识别异常信号；对于触发告警的事件，可进一步开展事件重构分析，还原事件发生过程并识别潜在原因。同时，基于聚合运行数据开展统计趋势分析，识别风险变化规律及潜在系统性问题，为后续风险评估提供依据。

分析过程采用“自动化筛查-人工复核”的协同机制：自动化筛查负责在大规模运行数据中高效识别疑似事件，人工复核则对识别结果进行确认与深度根因分析，最终输出结构化的风险分析结果和运行证据[6-7]。对于采用AI模型的ADS，可结合失效模式及后果分析、系统理论过程分析、场景化危害分析等方法，对运行阶段出现的异常行为、性能退化和边界场景事件开展分析，识别相关失效模式及风险来源，并为整改措施制定与测试场景库补充提供依据[1]。

对于具备人机协同特征的ADS，还需纳入人因相关运行指标作为补充分析维度[23]，例如接管请求后的响应时间、接管成功率、系统误报/漏报引发的用户异常行为及投诉等。这类指标反映驾驶员-系统交互过程中的行为风险与信任演化特征，可用于识别HMI设计缺陷、过度依赖风险及接管能力退化问题，并为安全策略优化与场景库扩展提供依据。

2.2.4　监测报告与交互

企业按照法规要求标准化编制并报送初始通知、短期报告与定期报告（图5），实施报告分级分类定向报送[6-7]，作为企业向监管侧传递运行安全证据的主要机制。根据事件性质和车辆活动范围，制造商需按时限要求，分别向型式批准主管部门、事件发生地及车辆运行地监管部门报送相关数据。

在数据保留方面，可建立分级保留策略，如原始数据在满足法律法规规定的留存期限后安全销毁，安全问题相关数据保留至整改完成，关键数据集保留用于长期趋势分析。

2.2.5　安全改进与验证

基于分析结果的安全改进，以软件算法更新、硬件优化、测试场景库扩充为主要手段。改进措施实施后，通过运行监测数据对其效果进行验证，确认风险是否得到有效控制，并为部署后安全状态评估与持续改进提供运行证据支撑。

2.3　ISMR与企业安全保障体系协同

通过运行阶段证据的持续收集与反馈，ISMR为企业安全保障体系提供运行证据支撑，推动多类安全工程体系协同运行，形成风险识别、分析、整改与持续改进的闭环管理机制。协同架构如图6所示。

（1）与功能安全及预期功能安全协同。利用ISMR收集运行阶段系统性失效、误触发和“未知-不安全”场景信息，为研发端提供运行阶段危害证据，用于动态修正危害分析与风险评估结果，并驱动测试场景库的迭代扩充[24-25]，从而提升安全设计对长尾风险的预见性与覆盖度。

（2）与网络安全协同。在统一的安全事件视图下，ISMR关联分析行驶安全异常与网络攻击行为，提

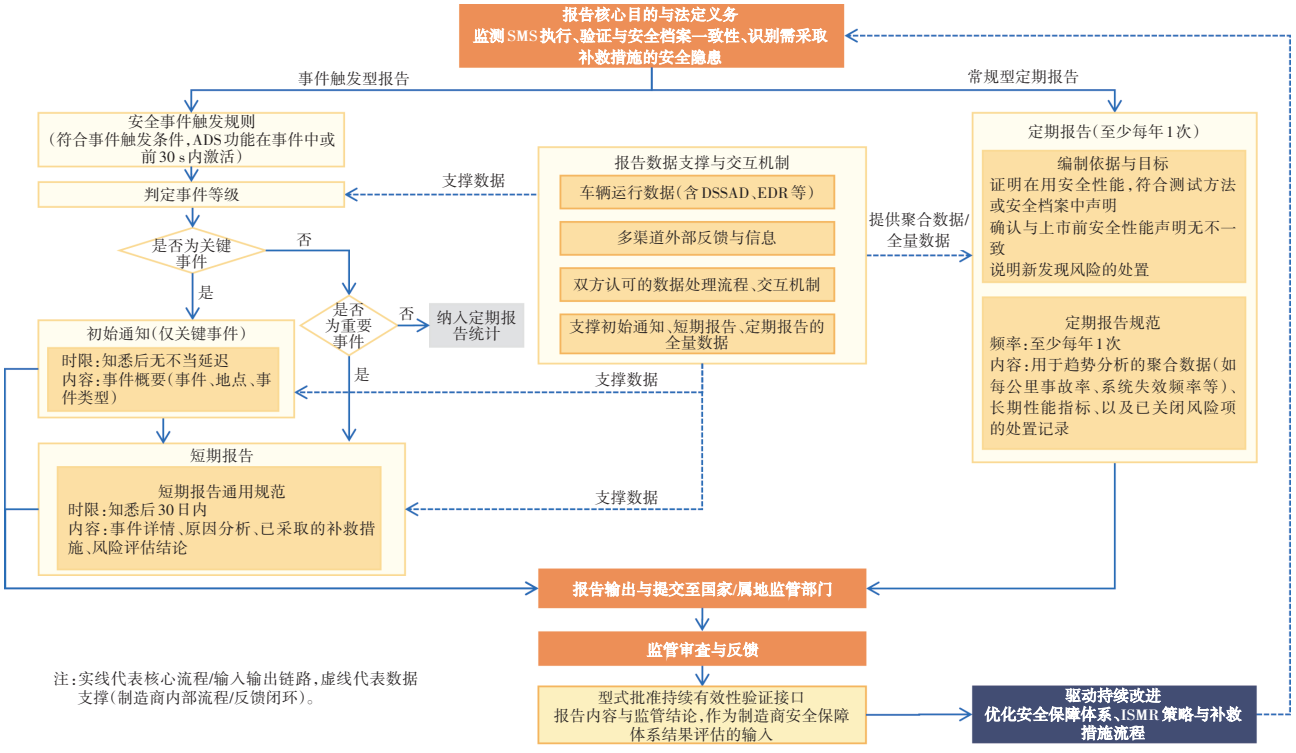


图5　ISMR运行阶段信息交互机制

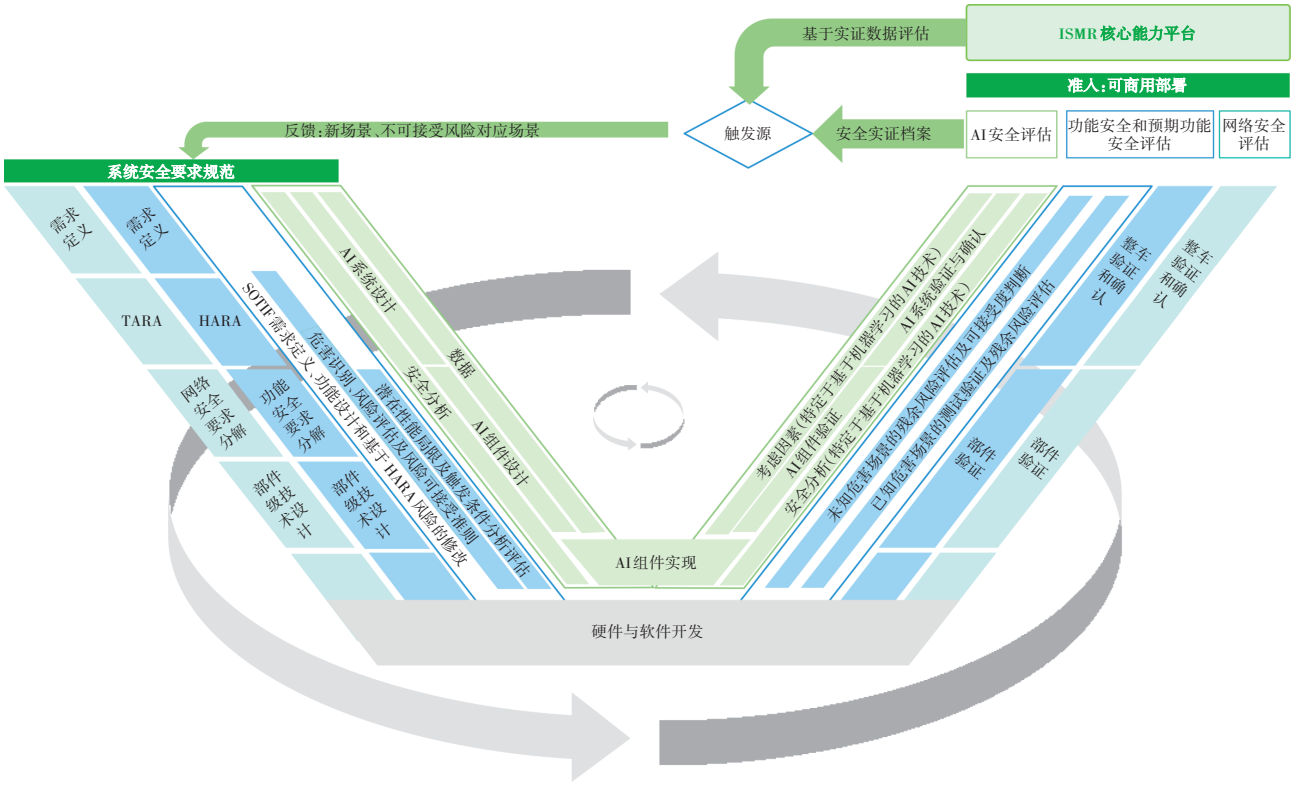


图6　ISMR与企业安全保障体系协同架构

升跨域安全态势感知与关联分析能力，支持快速判别事件成因（物理故障或网络攻击），并协调后续技术遏制与风险控制措施[4]。

（3）与软件更新管理协同。ISMR与软件更新管理流程协同[26]，形成更新-监测-验证的闭环，实现OTA升级过程的安全可控。基于OTA部署后运行数据，识别非预期退化或新风险，进而支撑扩大部署、回滚或追加修复等部署决策。

（4）与人工智能安全协同。ISMR从数据、证据与运行3个层面支撑AI安全治理：提供高质量运行数据

集用于模型训练和验证、输出合规与安全绩效证据用于安全档案维护，并支撑AI模型的现场数据收集与再验证活动[1]。

3 监管与企业协同治理机制

3.1 协同治理框架

企业实施体系为ISMR提供运行基础，管理部门与制造商之间的协同治理机制决定运行证据的生成、传递与反馈路径。本文构建“规则供给-运行执行-证据反馈”协同治理框架（图7），用以分析管理部门与制造商围绕运行安全证据的互动关系。

（1）规则供给层：由管理部门制定型式批准、ISMR、安全测试与评估、技术标准等规则，明确ADS在全生命周期内需满足的安全边界、评估要求与持续合规条件。

（2）运行执行层：由制造商在规则约束下，开展运行数据获取、风险分析、事件报告、安全改进与效果验证等活动，将制度要求转化为工程与管理实践，并持续生成反映系统运行状态的安全证据。

（3）证据反馈层：负责运行安全证据的汇聚、传递与再利用，支撑监管决策与制造商持续改进。管理部门基于多源证据开展风险评估、趋势研判、合规审查，并将评估结论与更新要求反馈至制造商，从而形成监管决策与制造商运行的动态联动。

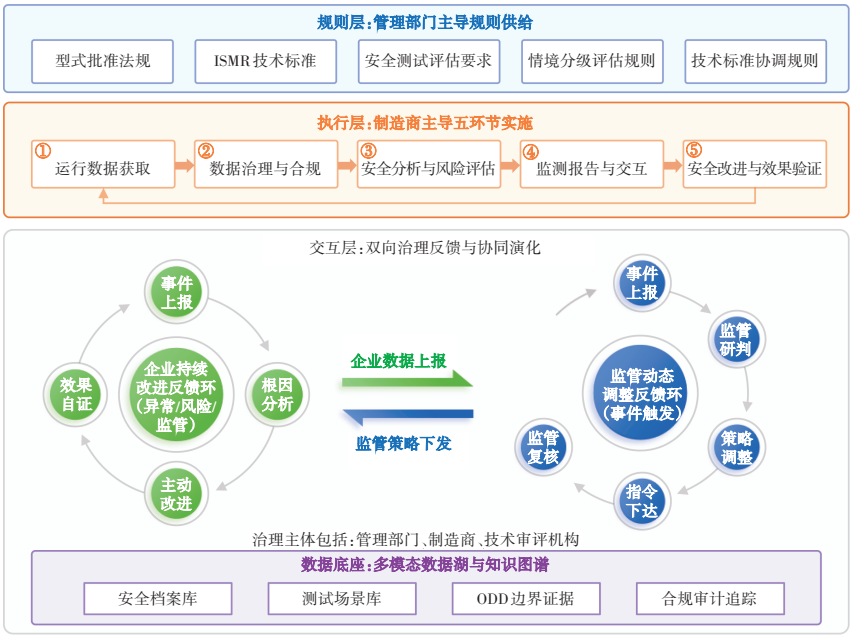


图7 规则供给-运行执行-证据反馈三层协同治理机制

3.2 协同治理双循环机制的运行机理

在协同治理框架下，系统运行表现为企业持续改进与监管动态调整两个相互耦合的反馈循环。其运行机理可分别从企业持续改进、监管动态调整2条路径进行分析。

3.2.1 企业持续改进反馈环

ISMR驱动的持续改进机制关注运行阶段安全边界的稳定性，即通过运行数据判断系统是否满足安全档案所定义的安全基线。制造商基于不低于法规要求的安全基线开展风险识别，主动触发改进措施，实施根因分析与安全优化，并通过运行证据持续验证改进效果，从而实现安全性能的动态维持。

标准化数据管道与统一数据结构是这一持续改进循环的基础条件。例如，欧盟L3Pilot项目通过统一数据格式与自动化数据质量工具[27]，实现多车型运行数据的标准化融合分析；我国智能网联汽车准入和上路通行试点形成的国家-地方-企业三级监测体系，也为ISMR持续改进机制提供了工程化实践基础。这些实践表明，数据治理能力是构建企业持续改进循环的关键前提。

3.2.2 监管动态调整反馈环

当发生碰撞、ODD边界外运行等特定安全事件并完成上报后，即进入监管动态调整反馈环。管理部门基于DSSAD、EDR等法定记录，对ADS行为与事件关系进行初步研判，并提出整改要求。企业落实整改后，再通过后续运行数据复核效果，验证风险是否已得到实质性控制。这一基于运行证据的反馈机制，将监管要求转化为以运行证据为约束条件的可验证治理路径。

在碰撞事件调查中，ISMR数据为ADS行为分析提供关键证据支持，用于辅助多源证据下的因果关系

认定，并支撑对事件是否属于合理可预见、可预防范围进行综合评估[6-7]。

4 事件触发的部署后安全评估框架

为形成完整的安全治理，本文基于ADS多支柱符合性评估理念[28]，构建事件触发的部署后安全评估框架（图8），由监测、分析与决策3个阶段构成。该框架以ISMR事件监测为触发端口，将碎片化的运行数据转化为结构化安全证据，用于支撑后续的安全评估与技术迭代约束。

4.1 监测触发

部署后安全评估触发机制可分为基于时间的常态评审触发、基于事件的法定触发以及基于变更的系统触发3类。其中，后2类与ISMR直接关联。

基于联合国WP.29、ISO/TS 5083：2025等法规标准对安全绩效指标的分类要求[6,8]，并结合试点运行暴露的风险问题与监管实践，本文构建如表2所示的ISMR安全绩效指标体系。该体系由滞后性指标、先导性指标、过程性指标、整改验证指标构成，用于支撑常态健康度监测、风险预警以及评估触发。

当安全绩效指标持续偏离运行基线，或出现重大性能异常与关键事件时，启动部署后安全评估流程。

4.2 根因分析、影响评估与风险重估

以触发信号为输入，本节对系统运行过程中的根因结构、影响范围及风险等级进行分析与重估。

（1）多维度根因定位：不同于传统交通事故鉴定，自动驾驶安全评估需要从系统行为、运行环境、数据与场景覆盖情况3个维度开展。首先，分析系统行为，包括事件发生前的规划轨迹、感知置信度及决策逻辑，以判断是否存在算法缺陷、功能异常或决策逻辑缺陷[1,3]。其次，评估运行环境，结合ODD与交通流数据，判断环境复杂度是否超出系统设计能力。最后，分析数据与场景覆盖情况，重点确认训练数据与运行场景之间是否存在分布偏移，判断问题来源于已知场景下的误判，还是运行过程中出现的新场景与未覆盖风险[1]。

（2）影响范围与风险重估：确定根因后，需从2个维度评估风险影响：一是车队波及分析，确认是单车个案，还是同型号、同软件版本的全车队共性问题。二是基于运行证据的动态风险重估。基于运行证据评估风险暴露水平、潜在后果严重程度及风险控制有效性，并与安全档案中的风险接受准则进行比较。若残余风险超出安全档案的可接受水平，应立即触发运行限制、功能降级等管控措施。鉴于AI系统的不确定性，重估过

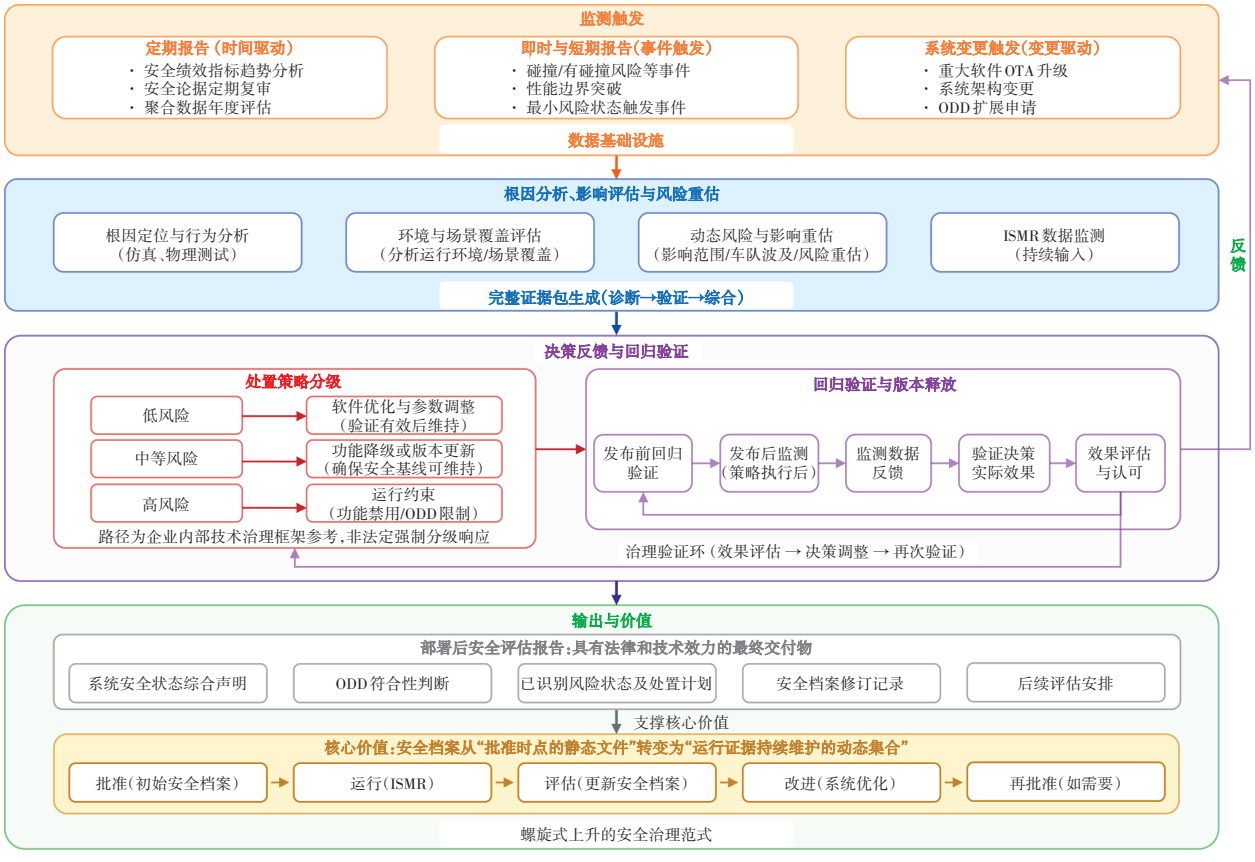


图8 基于事件监测触发与反馈的部署后安全评估框架

表2　ISMR安全绩效指标分类体系及典型指标[6,8-9,27]

指标类别	具体指标示例	典型阈值设定参考
滞后性指标(结果性度量,用于事后验证)	碰撞率、可归因事故率、责任碰撞频次	不高于安全档案定义的容忍上限或监管基线,或出现率趋势显著上升时触发调查
先导性指标(风险先兆,用于提前告警)	接管请求率、MRC触发频率、安全边界违反频率(如TTC阈值突破)、TTC<阈值事件的频率	基于安全档案承诺值设定初始阈值;运行稳定后,结合历史运行基线(如均值±2σ)进行动态调整,持续偏离即触发预警
过程性指标(运行阶段状态监测)	ODD外运行时长占比、传感器遮挡频次、算法推理时延超标率、关键功能不可用时长占比、数据质量达标率等	超过安全设计要求、安全档案约束值即触发报告,或作为早期信号进入调查流程
整改验证指标(闭环确认)	整改后同类场景紧急策略、风险缓解状态触发率、关联功能域的关键SPIs退化率、关键SPIs回归情况	整改后关键SPIs回归正常波动区间,相邻域无显著不安全退化(需结合工程判断与必要的统计分析)

注:本指标体系为示例性框架,基于试点阶段的监管需求、L3Pilot等项目经验设计。指标定义与阈值需结合具体ADS车型的安全档案、法规要求及运行数据基线进行具体设定,并根据运行数据动态调整。

程中需特别关注模型在相同输入下行为输出的风险分布波动情况[3]。

4.3　决策反馈与回归验证

(1)处置策略分级:根据风险评估结果,ISMR建立分级响应机制,将风险等级映射至差异化处置策略。低风险问题以软件优化与参数调整为主,并通过运行监测验证其有效性与风险变化情况;中等风险在确保安全基线可维持的前提下,可通过功能降级或版本更新实施控制;已判定存在不可接受风险的高风险情形,应优先实施运行约束措施(如功能禁用或ODD限制)以降低即时风险,并同步启动软件层根本原因分析与修复及版本迭代流程。该分级机制实现风险等级与OTA变更约束的统一映射,为运行阶段风险处置提供结构化决策依据[29]。

(2)回归验证与版本释放:任何风险处置措施在实施前,均需通过回归验证,以确认拟采取措施能够在目标触发场景中有效抑制不安全行为,并评估其对相关功能域及系统行为一致性的影响。回归验证通常结合仿真测试、场地测试、道路测试等方法,验证修复措施有效性,同时识别是否引入新的功能退化或非预期行为,从而避免局部修复导致系统性风险迁移。

在版本发布后,依托ISMR持续运行监测机制,对新版本在真实运行环境中的稳定性、风险控制有效性进行持续评估。最终通过修复版本发布与运行监测验证,实现风险控制、OTA更新约束与运行验证的闭环管理[3,29]。

5　研究结论

围绕ADS部署后安全治理,本文从制度分析与系统工程融合视角,研究ISMR的制度逻辑、实施体系及协同治理机制,主要结论如下:

(1)ISMR是支撑ADS型式批准从"准入阶段符合性验证"向"运行阶段持续评估"拓展的重要机制之一;ISMR与以算法性能优化为目标的数据闭环体系具有不同功能定位,前者服务于安全治理,后者侧重服务研发优化。两者在运行数据和工程支持体系层面协同互补,共同支撑ADS全生命周期安全保障。

(2)提出面向企业落地的ISMR实施机制,系统构建运行数据获取、数据治理、安全分析、监测报告与安全改进的治理闭环,并明确其与功能安全、预期功能安全、网络安全、软件更新、AI安全等体系的协同关系,为ISMR工程化实施提供方法框架。

(3)提出面向监管、企业协同的"规则供给-运行执行-证据反馈"协同治理框架,明确管理部门与制造商的职责分工,并通过企业持续改进、监管动态调整的双循环机制,形成基于运行证据驱动的协同治理范式。

(4)构建基于事件触发的部署后安全评估框架,通过监测触发机制启动评估流程,形成监测、分析与决策3阶段结构,实现对长尾场景、系统动态演化以及软件迭代风险的持续识别与闭环控制机制。

当前ISMR制度仍处于试点与规则建设阶段,相关分析主要基于公开政策法规、行业标准及工程实践经验。未来研究可围绕多源运行数据融合建模、安全绩效预测方法展开,并结合高保真场景重建与生成技术提升事件复现能力,同时探索安全档案动态更新与数实融合验证机制,以支撑更高可信度的自动驾驶系统安全治理体系。

参　考　文　献

[1] ISO. Road Vehicles-Safety and Artificial Intelligence: ISO/PAS 8800:2024[S]. Switzerland: ISO, 2024.

[2] FAVARO F, FRAADE-BLANAR L, SCHNELLE S, et al. Building a Credible Case for Safety: Waymo's Approach for the Determination of Absence of Unreasonable Risk[EB/

OL]. (2023-03)[2026-05-23]. www.waymo.com/safety.

[3] RAO A K, KELLER A J, KALRA N, et al. Challenges to the Monitoring of Deployed AI Systems(NIST AI 800-4) [R]. Gaithersburg: National Institute of Standards and Technology, 2026.

[4] United Nations Economic Commission for Europe. Uniform Provisions Concerning the Approval of Vehicles with Regard to Cybersecurity and Cybersecurity Management System: UN Regulation No. 155[S]. Geneva: UNECE, 2021.

[5] BUYSSE L, HABLI I, VANOOST D, et al. Safe Autonomous Systems in a Changing World: Operationalising Dynamic Safety Cases[J]. Safety Science, 2025, 191: 106965.

[6] UNECE. Proposal for a New United Nations Global Technical Regulation on Automated Driving Systems (ADS) [S/OL]. (2026-03) [2026-07-14]. https://unece. org/sites/default/files/2026-03/WP.29-198-07e.pdf.

[7] 工业和信息化部, 公安部, 住房和城乡建设部, 等. 关于开展智能网联汽车准入和上路通行试点工作的通知(工信部联通装〔2023〕217号)[EB/OL]. (2023-11-17)[2026-01-15]. https://www.miit.gov.cn/zwgk/zcwj/wjfb/tz/art/2023/art_be04a39345354b1ba4624262f86e7cfb.html.

[8] ISO. Road Vehicles-Safety for Automated Driving Systems-Design, Verification and Validation: ISO/TS 5083: 2025[S]. Switzerland: ISO, 2025.

[9] HILLER J, KOSKINEN S, BERTA R, et al. The L3Pilot Data Management Toolchain for a Level 3 Vehicle Automation Pilot[J]. Electronics, 2020, 9(5): 809.

[10] AVSC. AVSC Best Practice for Continuous Monitoring and Improvement after Deployment: AVSC00011202307[R]. Warrendale: SAE Industry Technologies Consortia, 2023.

[11] 贾贝贝, 陈贞, 李京泰, 等. 自动驾驶汽车法规与评估方法:研究进展与展望[J]. 汽车文摘, 2025(10): 19-23.

[12] 李睿恒, 王彩玉. 美国自动驾驶安全监管的联邦框架和地州实践[J]. 现代世界警察, 2025(11): 52-59.

[13] 工业和信息化部装备工业一司. 工业和信息化部等三部门召开智能网联汽车道路测试与示范应用工作会议[EB/OL]. (2026-04-14) [2026-05-23]. https://www.miit.gov.cn/jgsj/zbys/qcgy/art/2026/art_0a0efaeab3c84a3d9051ea1d17ee8737.html.

[14] LI X, WANG Z, HUANG Y, et al. A Survey on Self-Evolving Autonomous Driving: A Perspective on Data Closed-Loop Technology[J]. IEEE Transactions on Intelligent Vehicles, 2023, 8(11): 4613-4631.

[15] 黄岩军, 陈诗阳, 韦登伟, 等. 学习型自动驾驶决策算法闭环学习方法[J]. 汽车工程, 2026, 48(3): 491-501.

[16] 刘法旺, 徐晓庆, 陈贞, 等. 搭载自动驾驶功能的智能网联汽车安全测试与评估方法研究[J]. 汽车工程学报, 2022, 12(3): 221-227.

[17] 工业和信息化部. 中华人民共和国工业和信息化部公告2026年第1号(《道路机动车辆生产企业准入审查要求》和《道路机动车辆产品准入审查要求》)[EB/OL]. (2026-01-21) [2026-07-14]. https://www. miit. gov. cn/zwgk/zcwj/wjfb/gg/art/2026/art_38716cc92c6047169da8a3ec84fa940a.html.

[18] 国家互联网信息办公室, 国家发展和改革委员会, 工业和信息化部, 等. 汽车数据安全管理若干规定(试行)[EB/OL]. (2021-08-16) [2026-05-23]. https://www. gov. cn/zhengce/zhengceku/2021-09/12/content_564 0023.htm.

[19] 邱彬, 李广友, 薛晓卿, 等. 智能网联汽车数据交互与综合应用公共服务平台研究与构建[J]. 汽车工程学报, 2024, 14(5): 829-838.

[20] 宋琦, 武波, 刘永东. 关于智能网联汽车数据安全治理框架的探究[J]. 网络安全与数据治理, 2024, 43(3): 1-9.

[21] 李凤华, 李晖, 牛犇, 等. 车联网隐私保护技术研究[J]. 通信学报, 2023, 44(2): 1-18.

[22] 彭海洋, 计卫星, 刘法旺. 基于区块链的自动驾驶仿真测试数据存证模型[J]. 计算机应用, 2025, 45(8): 2421-2427.

[23] 贾萌宵, 吴亮, 赵豪杰, 等. 破解智能汽车"高潜力-低信任"研究综述:技术困境、组织障碍与环境约束[J]. 汽车文摘, 2025(10): 34-43.

[24] ISO. Road Vehicles-Safety of the Intended Functionality: ISO 21448: 2022[S]. Switzerland: ISO, 2022.

[25] ISO. Road Vehicles-Functional Safety: ISO 26262: 2018[S]. Switzerland: ISO, 2018.

[26] UNECE. Uniform Provisions Concerning the Approval of Vehicles with Regard to Software Update and Software Update Management System: UN Regulation No. 156[S]. Geneva: UNECE, 2022.

[27] HILLER J, SVANBERG E, KOSKINEN S, et al. The L3Pilot Common Data Format—Enabling Efficient Autonomous Driving Data Analysis[C]// 26th International Conference on the Enhanced Safety of Vehicle (ESV). Eindhoven, Netherlands: ESV, 2019.

[28] 刘法旺. 自动驾驶汽车安全测试评估方法[M]. 北京: 人民邮电出版社, 2024.

[29] 工业和信息化部, 市场监管总局. 工业和信息化部 市场监管总局关于进一步加强智能网联汽车产品准入、召回及软件在线升级管理的通知[EB/OL]. (2025-02-28) [2026-02-04]. http://www. gov. cn/zhengce/zhengceku/2025-02/28/content_7000000.htm.

（责任编辑　**明慧**）